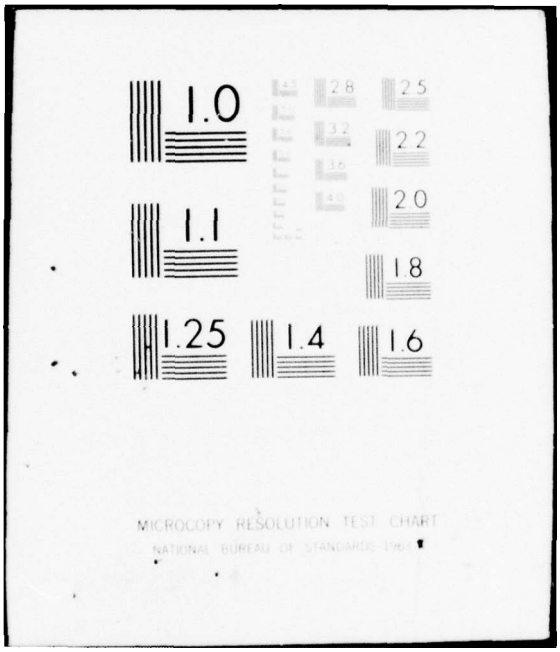
AD-A034 287 CHIEF OF NAVAL EDUCATION AND TRAINING SUPPORT PENSAC--ETC F/G 5/9
THE RELATIONSHIP OF READABILITY AND CONCEPTUAL LEVEL TO PERFORM--ETC(U)
OCT 76 R J BIERSNER, R E DOUCETTE

UNCLASSIFIED CNETS-3-76 NL

1 OF 1
AD A034287

END
DATE FILMED
2-77


1.0
1.1
1.25
1.4
1.6
2.8
2.5
3.2
2.2
3.6
2.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963

CNETS REPORT 3-76

THE RELATIONSHIP OF READABILITY AND CONCEPTUAL LEVEL TO PERFORMANCE ON A NAVY NON-RESIDENT CAREER COURSE

OCTOBER 1976

DEPARTMENT OF THE NAVY
CHIEF OF NAVAL EDUCATION AND TRAINING SUPPORT
PENSACOLA, FLORIDA 32509

IN REPLY REFER TO

Code 01A2:tdh
1552

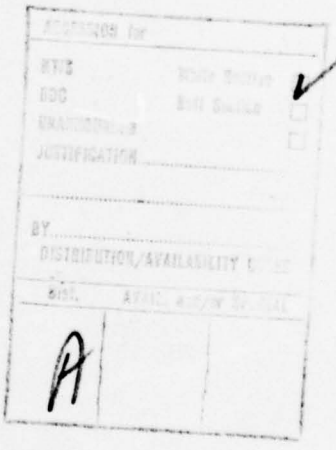
NOV 18 1976

1. CNETS Study Report 3-76, "The Relationship of Readability and Conceptual Level to Performance on a Navy Non-Resident Career Course" is promulgated for information.

2. The conclusions and recommendations contained in the report are those of the writer and are not necessarily those of the Chief of Naval Education and Training Support.

3. This publication has been reviewed under the provisions of SECNAVINST 5600.16 and is approved.

WAYNE P. HUGHES, JR.



CNETS REPORT 3-76

THE RELATIONSHIP OF READABILITY AND CONCEPTUAL LEVEL TO PERFORMANCE ON A NAVY NON-RESIDENT CAREER COURSE

(CNETS Field Task Assignment Number 60044)

PREPARED FOR

THE CHIEF OF NAVAL EDUCATION AND TRAINING SUPPORT

BY

ROBERT J. BIERSNER, Ph.D.
LCDR USN MSC

AND

ROBERT E. DOUCETTE, B.S.[1]

October 1976

TABLE OF CONTENTS

1.0 SUMMARY. The relationships of readability (as measured by the Flesch and Kincaid formulas) and conceptual level (CL) to performance on study questions contained in the Non-Resident Career Course (NRCC) Hull Maintenance Technician 3&2, NAVEDTRA 91510 were analyzed. Two samples of study questions, a nonerror sample and an error sample, were drawn from this NRCC. The results show that readability between the nonerror and the error samples, or between the textual passages associated with these samples, does not differ significantly. The results do show, however, that the conceptual level (thinking) required to answer the study questions is related to error rate. That is, high error rate study questions are associated with higher conceptual levels than are zero error rate study questions. Methods are presented for using high and low CL study questions to improve the training effectiveness of NRCCs. The recommendation is made that retention of high and low conceptual level study questions be analyzed because previous research has indicated that better learning and memory may be associated with study questions that are highly conceptual. The recommendation is also made that the relationship between NRCC performance and verbal intelligence (measured through the General Classification Test) be determined because verbal intelligence is most likely involved in performing well on highly conceptual study questions. In addition, readability formulas should be developed to include conceptual level as a factor in determining reading grade levels. The inclusion of a conceptual level factor in readability formulas could improve the validity of these formulas.

2.0 ACKNOWLEDGMENTS. This task could not have been accomplished without the cooperation of the CNETS/NETPDC Reading Skills Measurement Working Group, which was formed under CNETS Field Task Assignment No. 60044. Specific contributions by members of this group include: the technical advice and recommendations of Dr. Margaret Smith; the data collection procedures and organization provided by Mr. Andrew Dow and Mr. Robert Pallme, and the insight afforded by Mr. Pallme that led to the eventual development of the 4-point conceptual level rating scale. Appreciation is also extended to OSCM Maurice McCoy and Mrs. Carol Jernigan who collected the readability data, and to Miss Donna Airy and Mr. Michael Pantle who served as judges in determining the conceptual levels of the NRCC materials. The programming support provided by Miss A. Lucille Shirk was invaluable and greatly appreciated.

3.0 PURPOSE. This report presents a preliminary analysis of the relationship between performance on study questions which were administered routinely in conjunction with a current Navy Non-Resident Career Course (NRCC) and the readability of the Rate Training Manual (RTM) passages from which these study questions were taken. In addition, the association between study question performance and the readability of the study questions will be determined, as well as the relationship between study question performance and the conceptual level required to answer the questions correctly using the information provided in the associated RTM passage. The data and interpretations found in this report may be useful to the Chief of Naval Education and Training Support and others within and outside the Naval Education and Training Command in developing and implementing policies and procedures to make reading materials and nonresidential instruction more useful to the consumer.

4.0 BACKGROUND. A number of recent reports (references 1, 2, 3, and 4) have demonstrated that a substantial difference may exist between the readability of Navy written materials and the reading skills of the Navy personnel who are required to read these materials in order to perform well or to advance in rate. Readability of written materials is measured by a variety of formulas which are highly intercorrelated and which are purported to measure the level of reading difficulty in terms of the average grade level of education which would be required to comprehend the textual information. This average education level is referred to as Reading Grade Level (RGL).

In response to the above findings, the recommendation has been made (OP-099 speedletter 9911/550 of 13 March 1974) that the RGLs of Navy training and technical materials should be reduced in order to improve occupational performance and advancement opportunities among Naval personnel. Later recommendations (reference 1) cautioned against arbitrary lowering of RGLs, and stated that while the formulas which are used to measure RGLs may be useful in monitoring the overall reading difficulty of written materials, the validity of these formulas in measuring accurately the reading comprehension of Navy enlisted personnel for Navy technical and training materials has not been established satisfactorily. (An extensive discussion of the validity problems associated with these formulas is provided in reference 5.) Inasmuch as most of these formulas determine RGLs using some combination of average sentence length, number of syllables per word, and the frequency with which words are used in common forms of writing, the possibility exists that writers could use shorter sentences, reduce the complexity of words, and use more frequent or common words and still write material which is difficult to comprehend. In addition, little has been done either among military or civilian organizations to show that conscientious use of these parameters (shorter sentences, less complex words, and more common words) results in improved comprehension or performance. Although the Air Force (reference 6) has demonstrated that reductions in the Automated Readability Index (ARI) for a technical manual significantly improved performance on an associated multiple-choice test, these findings must be interpreted with caution. The absence of pre- and post-revision samples, and the failure to have experienced

editorial writers compare content and style characteristics between the two manuals, leave open the possibility that modifications in content and style (in addition to reduced RGLs) may account for the performance improvements which were observed. The advice found in professional writing guides (references 7 and 8), as well as the recommendations made by researchers (references 1 and 9), would indicate that RGLs, while important to total reading comprehension, should not be emphasized to the exclusion of content and style.

A problem that was ignored in the above Air Force research, and that is often ignored by writers as well as those who are responsible for using written materials in training, is the relationship between test questions and the textual material which provides the information necessary to answer these questions. Little is known about the relationship of item error rate to (a) the readability (RGLs) of the questions, and (b) the extent to which questions duplicate the textual material (i. e., differences in error rates between verbatim questions and those questions which require the reader to conceptually organize or think about the written information in order to provide a correct answer). The possibility exists that comprehension, as measured by test questions, is not related solely to the readability and style of the textual material, but is related as well to the readability of the test questions and to the conceptual level required to answer the questions correctly. Previous findings (references 9 and 10) have indicated that the study question technique used in Navy NRCC courses, in which the test questions (referred to as "study questions") are used to guide a reader through the NTMs, may not be training effective.

This technique may be ineffective for retention of the textual information which answers the study questions and for learning textual information which is incidental to these questions. The results found in reference 10 show further that error rates among study questions appear to be related to the extent to which the reader must (a) search for the answer in the RTM, and (b) integrate several different types of information in answering the question. The findings in reference 9, however, may not be valid for Navy use because of the small sample size, the high reading skill level of the subjects, and inadequate research design and statistics, while reference 10 used highly subjective techniques to judge the textual factors that were related to error rates among the study questions.

The following data were collected in order to determine if error rates for NRCC study questions are related significantly to (a) the readability (RGLs) of the textual materials (RTMs) which contain the answers to these questions, (b) the RGLs of the study questions, and (c) the extent to which the RTM information must be integrated and organized conceptually by the reader in order to answer the study questions.

5.0 METHODS. The following subjects, procedures, and statistical techniques were used in collecting and analyzing the present data.

5.1 Subjects. The subject group consisted of 30 Navy personnel[2] who submitted answer sheets in order to obtain credit for the NRCC entitled "Hull Maintenance Technician 3&2" (HMT 3&2). The HMT 3&2 course was used because (a) this course was judged by several experienced Navy writers to be well written[3], (b) the HT (Hull

Technician) rating was judged to be a fairly representative Navy rating[4], and (c) much of the data had already been submitted and analyzed in conjunction with another NPRDC research project.

The HMT 3&2 course is taken routinely by apprentices (or "strikers") who are trying to qualify for the HT3 (Hull Technician Third Class) advancement in rate examination, or by those personnel who are rated already as HT3s, but who are trying to advance to HT2 (Hull Technician Second Class). The present sample is not necessarily representative of the total HT Apprentice and HT3 populations enrolled in the HMT 3&2 course, but is more likely to be representative of those personnel who completed the course and whose commands (ships or stations) were diligent and cooperative enough to return the completed answer sheets. The present sample therefore probably represents a more intelligent and motivated group than those who did not complete the course or who were assigned to commands which neglected to return answer sheets. In addition, the present subject group was taken from among those who first submitted answer sheets, so the possibility exists also that they may not be representative of the total group which eventually submitted these sheets.

Although aptitude and career data were not available for the present subjects, a review was made of these data for HT Apprentices and HT3s who had passed the HMT 3&2 course successfully and who took the examinations for advancement to the next higher rate (HT3 and HT2) in August of 1975. These data included average GCT, the percentage of personnel at the HT Apprentice and HT3 levels who had been at these

rate levels 12 months or less (months in rate), and the percentage of personnel at each of these rate levels who had been in the Navy 36 months or less (length of service). In order to determine whether or not MT is a representative Navy rating, the above aptitude and career data for MT Apprentices and MT3s were compared to similar data for personnel taking any August 1975 examination for advancement to comparable rate levels (Overall Apprentice and Overall Petty Officer 3 groups).

5.2 Materials. The MMT3&2 course, similar to most NRCCs, consists of an assignment booklet and an RTM. Each assignment in the assignment booklet covers 1 or more subject matter areas in Full Maintenance Technology. The chapters in the RTM discuss in detail the subject matter areas presented in the assignment booklet. Each assignment consists mostly of multiple-choice study questions and a few true-false and matching study questions. Previous research (reference 1) has shown that the RTM for MTM3&2 has an overall Flesch RGL between the 10th and 12th grades and a Kincaid RGL of 10.72 grades. The MTM 3&2 assignment booklet has an overall Flesch RGL between the 10th-12th grades.[5]

Both the RTM and assignment booklet are issued together, and the course is completed at the convenience of the student during off-duty hours with minimal supervision. The student is told to study the specific pages of the RTM, which are identified at the beginning of each assignment, and then answer the associated study questions in the assignment booklet. As demonstrated in reference 10, however, the reverse procedure appears to occur more often--the student analyzes the questions and then tries to find the answers in the RTM. This procedure probably

makes for sporadic and mixed coverage of the textual material contained in the RTMs. Responses to the multiple-choice study questions are entered on a separate answer sheet by erasing 1 of 4 columns (2 columns in the case of true-false questions; 2,3, or 4 columns for matching questions). If an incorrect answer is entered (erased), the incorrectly erased column refers the student to the RTM page which begins the topical discussion from which the correct answer may be derived. The student then reads this material and enters another response. The appropriate RTM material is not too difficult to locate because the questions usually cover sequential RTM material, a condition which probably encourages the student to refer to the questions before reading the text. Underlying the answer sheet is a duplication sheet which provides a permanent record of the erased responses so that a determination can be made of the number of responses entered before the correct answer was chosen.

For the purpose of the present analysis, those study questions which had 2 or more erasures were counted as incorrect, while those questions which had a single erasure were counted as correct. Only the study questions contained in the first 12 (of 27) assignments were analyzed in order to reduce the data collection and statistical requirements. In addition, only multiple-choice and true-false study questions (and the RTM passages associated with these questions) were included in the analysis because the readability of matching questions was too cumbersome and unreliable to determine. In order to further maximize the differences between the study questions, and to improve the classification of the study questions into error and nonerror samples,

only those questions which were not missed by any of the subjects (nonerror sample) and those questions which were missed by at least 5 of the subjects (error sample) were used in the present analysis. This differentiation resulted in 106 questions in the error sample and 77 questions in the nonerror sample. In order to make reliable judgements of the reading grade levels (see section 5.3 below), those study questions that referred to diagrams in the RTM were also removed from the 2 samples. As a result of this procedure, 7 study questions were removed from the error sample and 3 study questions were removed from the nonerror sample. The answers to the remaining questions in the error sample and the nonerror sample could, therefore, be referenced directly to a specific and delimited passage of textual material within the RTM. These textual passages were provided by the writers of the HMT 3 & 2 course.

In an additional sampling procedure, study questions were removed from further analysis if either the study question or the associated RTM passage did not conform to the readability parameters established by Flesch (reference 11).[6] This procedure resulted in disregarding 19 questions from the error sample and 9 questions from the nonerror sample. Following this procedure, the total number of study questions in the error sample was 80, while 65 study questions remained in the nonerror sample.

5.3 Measuring RGLs and Conceptual Levels. The Reading Grade Levels (RGLs) of the RTM passages described above, as well as the RGLs of the associated study questions, were determined using both the Flesch Reading Ease Score (the Flesch formula) and the correction to the Flesch

formula which was developed recently by Kincaid (the Kincaid formula). The procedures involved in applying these formulas to this type of NRCC material, which are too extensive to describe in detail in this report, have been presented fully in an earlier CNETS publication (reference 1). Described briefly, however, the Flesch formula uses the number of syllables, words, and sentences in textual samples of at least 100 words to determine RGL. In the case of those RTM passages and study questions which consisted of less than 100 words, the RGLs were determined using correction procedures recommended by Flesch (reference 11). The Kincaid formula modifies the resulting Flesch RGL by correcting for the reading skills which Navy personnel have demonstrated for this type of Navy training material. The Kincaid formula has been found to reduce the majority of Flesch RGLs for NRCCs between 1 and 2 grade levels (reference 1).

The Conceptual Level (CL) required by the students to answer the study questions from the textual information provided in the associated RTM passages was determined by 2 independent judges. The judges were psychology students (a female and a male) who were seniors at a local university. Both judges were about 25 years of age and were earning above average college grades. Neither judge was familiar with NRCC writing or testing procedures, and neither judge had experience with RTM terminology or procedures.

The 2 judges were asked to use a 4-point rating scale in determining the extent to which NRCC students would have to think about the information which was provided in the associated RTM passages in answering each of the 145 total study questions (80 questions in the

error sample and 65 questions in the nonerror sample). The 4-point rating scale was adapted from the first 4 cognitive domain levels developed by Bloom (reference 12). The judges first read the question stem and the 4 choices (or, in the case of true-false items, the item statement), and then read the associated NTM passage before making each judgment. They were allowed to re-read the passage, statement, question stem, and choices as often as they wanted before making a judgement. They were not told, however, which choice was correct before rating the study question. The rating scale, ranging from a score of 1 (lowest CL--little thinking required to answer the study question) to 4 (highest CL--much thinking and organization required to answer the study question), is described in Appendix A.

The 4-point scale was explained carefully to the judges at the start of a training session, and they were then allowed to practice on study questions taken from later NTM3&2 assignments (assignments 13 to 27). These practice study questions consisted of an equal number of error study questions (those questions missed by at least 5 of the subjects) and nonerror study questions (those questions which were not missed by any subject). The error and nonerror questions were arranged in random order. After rating each practice question, the judges explained the logic they used in making these ratings. They continued practicing until they demonstrated consistency in separately scoring 14 out of 15 practice questions (this criterion was reached after 2 practice trials). The 2 judges then rated the 145 study questions for 3 hours a day over the next 5 days. The judges also scored 5 practice questions at the beginning of each of these 5 days in order

to maintain familiarity with the rating procedures. The judges were free to use a dictionary while reading the MTM passages and study questions, and to consult with MTM writers and subject matter experts during this period (neither of which they did). The judges did not communicate with each other while making the judgements, nor were they told which study questions were in the error and nonerror samples. After rating each study question, the judges were told the correct choice, as well as the ratings they each had made of that study question. If these ratings differed, they were asked to explain the logic which they had used in making the judgements. The original judgements, however, were used for calculating the present results and were not modified after these explanations.

5.4 Statistical Procedures. Reliability between the 2 judges in rating CLs was determined using an intraclass correlation. In order to insure that these separate CL judgements not only varied in a consistent fashion, but were truly similar, a t-test for independent samples was used to compare the means of the ratings made by each judge. Differences in RCLs and CLs for the error and nonerror samples were also tested for significance using t-tests for independent samples. (The CL score for each study question within the error and nonerror samples was the average of the ratings assigned to that question by the 2 judges.) The t-test for independent samples was used as well for comparing the GCT scores of MT Apprentices and MT3s with Overall Apprentice and Petty Officer 3 groups, while the chi-square test was used to compare differences in career data (length of service and months in rate)among these groups. For the GCT data, variances were found to differ significantly for the MT Apprentice

and HT3 groups on comparison to the Overall Apprentice and PO3 groups. Consequently, the Cochran and Cox method (reference 13) was used to compute the t-tests involving GCT means. Levels of significance for all tests were $p \leq .05$ (two-tailed).

6.0 RESULTS. The following sections present the findings for comparisons between HTs and other ratings, differences in PCLs between the error and nonerror samples, reliability of the judges in determining the CLs of the study questions, and differences in CLs between the error and nonerror samples.

6.1 Comparisons Between HTs and Other Ratings. Table 1 presents the GCT and career data for HT Apprentices and HT3s, as well as GCT and career data for those who were in the Overall Apprentice and Petty Officer 3 groups. This table shows that significantly more personnel had been at the HT Apprentice and HT3 rate levels 12 months or less than personnel at comparable rate levels in the Overall Apprentice and Petty Officer 3 groups. In addition, significantly more HT3s had been in the Navy 36 months or less than personnel in the Overall Petty Officer 3 group. The proportion of HT Apprentices and Overall Apprentices who had been in the Navy 36 months or less did not differ significantly. Table 1 also provides evidence that both HT groups were significantly lower in verbal intelligence (GCT scores) than the Overall Apprentice and Petty Officer 3 groups. Inasmuch as these data show that HT Apprentices and HT3s are not representative of other (overall) comparable Navy ratings in verbal intelligence and months in rate, and that HT3s are not representative of other Petty Officer 3s in length of service, the results described below (sections 6.2 and 6.4) may not be replicated

for personnel in other Navy ratings. A further assumption, which remains unsubstantiated, is that the subjects used in the present analysis would be comparable to these HT Apprentices and HT3s in GCT scores, months in rate, and length of service.

6.2 Differences in RGLs Between Error and Nonerror Samples. The mean RGLs resulting from the Flesch and Kincaid formulas for the study questions and associated RTM passages in the error and nonerror samples are shown in Table 2. The t-tests presented in this table show that differences in Kincaid and Flesch RGLs between the 2 samples were not significant, either for the RTM passages or the study questions. (A noteworthy finding is that RGLs for the error sample were lower than RGLs for the nonerror sample in most cases, although these differences did not reach significance.) Differences in error rates between the 2 samples are unlikely therefore to be attributed to the readability of either the RTM passages or study questions. The average RGLs for the RTM passages and study questions in the error and nonerror samples also correspond closely to the average RGLs found for the entire HMT3&2 assignment booklet and RTM in previous research (reference 1).

6.3 Reliability of CL Judgents. The 2 judges were found to be in high agreement in rating the CLs of the 145 study questions. The intraclass reliability between the 2 judges for the 145 ratings was 0.836. The total sample of 145 study questions and associated RTM passages were assigned an average CL rating of 2.36 (sd = 0.91) by the first judge, and an average CL rating of 2.32 (sd = 0.93) by the second judge. The appropriate t-test showed that these means did not differ significantly ($t = 0.369$; $df = 288$; p = ns). Although these results do not substantiate

the validity of these judgements, this level of agreement in ratings indicates that under these conditions, the 3 judges were highly consistent in applying the standards which were used to rate the CLs of these study questions. Examples of study questions and associated PTM passages which were rated by the judges at the 4 conceptual levels are presented in Appendix F.

6.4 Differences in CLs Between Error and Nonerror Samples. The mean CL score for the error sample was 2.53 (sd = 0.88), while the mean CL score for the nonerror sample was 2.11 (sd = 0.76). The t-test resulting from a comparison of these 2 means was found to be significant ($t = 3.06$; $df = 143$; $p < .005$). These findings show that the average CL score for the error sample was significantly higher than the average CL score for the nonerror sample. This finding indicates that conceptual difficulty is involved significantly in the different error rates between the 2 samples.

7.0 DISCUSSION. The above results show that readability, as measured by the Flesch and Kincaid formulas, has little to do with performance (error rates) on the NRCC for IMT3&2 which is administered under these conditions. These findings appear to demonstrate that stylistic writing factors, such as those factors represented by CL scores, may be more important than readability to this type of performance in these training situations. Before discussing these PGL and CL effects further, however, the present results should be interpreted with caution because (a) the size of the subject group was small ($N = 30$) and selective (chosen from among those personnel who both completed and returned the IMT3&2 course), (b) the above GCT and career data indicate that IT Apprentices

and MT3s (the populations from which the present subjects are assumed to be drawn) are not entirely representative of other Navy ratings, (c) the MMT3&2 course may not be representative of other NPCCs inasmuch as these MMT3&2 course materials are recognized by professional NRCC writers as being especially well written, (d) the validity of CL ratings has yet to be demonstrated fully, (e) although the present findings indicate that CL ratings may be involved significantly in NPCC performance, this conclusion must await cross-validation of the present results using different judges, and (f) measures of readability other than the Flesch and Kincaid formulas (especially measures which involve comprehension of this type of course material among those specific Navy personnel who must read these materials in order to qualify for rate advancement) may yet demonstrate a relationship between readability and NPCC performance. In addition, the importance of conceptual level indicates that measures of verbal intelligence (such as GCT) may be related to NPCC performance, with those personnel who have higher GCT scores performing better than those with lower GCT scores. This effect may be most apparent for questions in the error sample which have high CL scores. This effect could not be analyzed in the present situation, however, because GCT scores were not readily available for this subject group. The above precautions indicate that (a) the present findings should be replicated on a much larger and more representative subject group (from a variety of Navy ratings and rate levels), (b) additional NPCCs should be used, (c) CL ratings should be replicated and validated, (d) replication should involve readability measures other than the Flesch and Kincaid formulas, and (e) a determination

should be made of the relationship between poor NRCC performance (high error rates) and verbal intelligence.

The present findings appear to reemphasize the basic deficiency in most readability formulas--the failure to account extensively and accurately for reading comprehension. These results show that formulas which manipulate sentences, words, and syllables, even if corrected for the reading skills which Navy personnel have demonstrated for related material, cannot measure the fine and subtle conceptual relationships which appear to exist between textual material and criteria of comprehension and performance effectiveness (such as error rates for questions). These formulas can account only for that gross level of comprehension (or conceptual) difficulty which is related to sentence, word, and syllable complexity, and can do little to differentiate the conceptual level required to comprehend material which is similar in sentence, word, and syllable structure. The present analysis overcame this deficiency by having judges who were naive about the subject matter (but who were not necessarily naive about study question difficulty) rate the conceptual level which was required to comprehend the textual material (RTMs) sufficiently well to perform effectively on the study questions. These results indicate that readability consists of more than sentence length, word frequency, and word complexity, but consists also of the cognitive processes involved in organizing and interpreting textual material. Although the reading skill which measures these cognitive processes is referred to as "comprehension", most reading experts have failed to specify comprehension in terms of the relationship which exists between textual material and performance criteria, while

none of the available reading formulas incorporates a separate comprehension or CI factor into computations of readability. The Cloze technique probably accounts for comprehension better than any other available reading formula, but the Cloze score may be confounded by vocabulary development and reading speed. This deficiency most likely results from the dependent relationship which Cloze scores have with standard measures of reading skill (reference 4). The validity problems associated with standard tests of reading skill further confound interpretation of Cloze scores.

Many of the above criticisms of existing readability formulas, especially the Cloze technique, have been made previously in a report by Carver (reference 9). In this report, Carver also developed a new readability formula, the Rauding Scale, which appears to correct many of the deficiencies associated with the earlier formulas by accounting for a comprehension factor. Similar to the CI rating method used in the present analysis, the Rauding technique uses personnel who are familiar with reading skill development (experienced teachers or graduate students enrolled in reading training courses), and who have passed a screening test (the Rauding Scale Qualification Test), to judge the readability of textual material. These judges appear to use both the input difficulty (sentence, word, and syllable complexity) and storage difficulty (the difficulty in organizing and integrating the input factors) in determining the readability of textual materials. Although the validity of an independent comprehension factor within the Rauding Scale remains to be demonstrated, this technique nonetheless warrants serious attention by writers and managers of military training materials. The

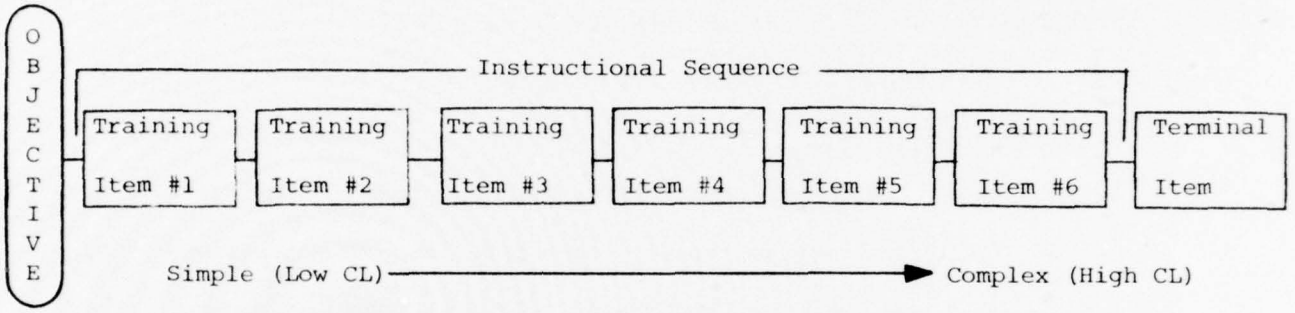
application of Rauding Scale techniques to the present NRCC data would appear to be an excellent means of validating and establishing the usefulness of the Rauding Scale in measuring the readability of military training materials.

Inasmuch as the above results indicate that error rates in NRCC courses are not related to the readability of either the PTM passages or the study questions, but are associated instead with the conceptual level required to answer the study questions correctly from the information which is provided in the associated PTM passages, a problem arises as to whether or not the conceptual levels of the difficult study questions should be made easier in order to reduce the error rate. Two lines of evidence indicate that although lowering the conceptual level may reduce error rates, such a procedure may impair overall training effectiveness. First, if the assumption is valid that the structure of NRCCs (especially the verbation characteristic of nonerror--low CL--study questions) permits students to read only a fraction of the RTMs (reference 10), then the more analytical and integrative characteristics of the high CL study questions would probably result in more thorough reading of the RTMs in order to obtain the information necessary to answer these study questions correctly. The effect of this more extensive reading should be improved learning of the textual material, especially improved learning of material that is incidental to answering the study questions correctly. Secondly, some evidence (reference 14) shows that textual material that is associated with high CL test items is remembered better (after 7 days) than textual material that is related to low CL (verbatim) test items. This retention effect

was found to be most significant for training situations in which the high CL test items were administered in a pre-test fashion, which is similar to the study question technique used in NRCCs. High CL test items (study questions) appear to result in better structured and organized long-term memory than low CL test items (study questions). These data demonstrate that while high CL study questions may be more difficult to learn initially than low CL questions, the improved retention of high CL questions once learned may make such questions more training effective than low CL questions. The use of high CL (high error) study questions in NRCCs should continue therefore because these types of questions appear to make NRCC students read more of the RTM material and remember the material which they do read better than if these questions were not included in the course materials. Perhaps the training effectiveness of NRCCs could be improved substantially if more high CL study questions were included in the present courses.

The order or distribution of high and low CL study questions within an assignment may also have important training consequences. Although each of the present NRCC assignments specifies several instructional objectives, the written material associated with these objectives (study questions and RTMs) do not appear to progress in a systematic fashion from simple to the more complex levels within the cognitive domain developed by Bloom (reference 12). At present, these written materials appear to be associated randomly with each objective. Perhaps NRCCs would be more training effective if these written materials were organized initially around basic knowledge factors (as indicated by study questions with low CL ratings) and progressed systematically

to the more complex cognitive factors of application and analysis (as indicated by study questions with high CL ratings). This method would require that (a) the training objective be stated in measurable terms, (b) the concept be presented in a series of training items, and (c) performance (application, analysis, and so forth) be tested in a terminal item. For example:



The CL rating technique described in this report could be used to assist NRCC writers in rearranging and reorganizing the written materials associated with instructional objectives. In this way, NRCC students may be better prepared to understand and answer correctly the more difficult study questions, as well as have better retention of both simple and complex information for longer periods than may presently be the case. The improved training effectiveness of NRCCs promised by this rating technique, as well as the ease of applying the technique to present NRCCs, should warrant an evaluation of the technique in the near furture.

The CL ratings described in this report may also have another application in the management of NRCCs. In reviewing the data from the error sample, about 47.5 percent of the study questions in this sample were found to have a CL rating of 2 or less. This finding indicates that these study questions required little thinking to answer, but were nonetheless frequently missed. The judges, as well as the authors of this report, perceived these study questions to be poorly written--ambiguous, misleading, or deceptive ("tricky"), or containing faulty technical information. Using samples of regular NRCC students in the future, comparisons could be made between the error rates and the CL ratings for study questions in an NRCC. Those study questions with high error rates and low CL ratings could then be identified and reworded, repositioned, or eliminated. As demonstrated in this report, this technique would be inexpensive because judges can be trained easily and, with little experience, a judge can rate reliably 10 to 20 study questions an hour.

Another application of this rating technique would be to judge the conceptual difficulty of information contained in graphs and figures. At present, a reliable technique or formula similar to those used to measure the readability of textual material does not exist for graphic materials. The use of judges and a rating scale as described in this report could provide a measure of the conceptual processing that may be necessary to understand the information contained in graphic material, or in combinations of graphic and textual material.

The above evidence indicates that more research should be done to measure retention of NRCC information over long periods, and that differences

in retention between information contained in high and low CL study questions should be analyzed. This research should also document whether or not more extensive reading of RTMs (as demonstrated by incidental learning) occurs in association with high CL than low CL study questions. An analysis should be made as well of the training effectiveness (learning and retention of study questions) resulting from rearranging and reorganizing (but not rewriting) the written materials (RTMs and study questions) associated with NRCC instructional objectives. Such rearrangement and reorganization would occur in the progressive, systematic (simple to complex) fashion described above according to the CL ratings which have been assigned by judges to the study questions that are used to support the instructional objectives. Inasmuch as NRCCs are a primary training medium for the Navy Reserve, and are also an important ingredient in the advancement of Navy enlisted personnel, these findings would have implications for the preparedness of the Navy Reserve, as well as the validity and management of the advancement in rate process.

8.0 CONCLUSIONS. The following conclusions are derived from the above findings:

8.1 NRCC performance, as represented by error rate on study questions, does not appear to be related to the readability level of NRCC study questions or associated RTM passages as determined by a popular, conventional measure of readability (the Flesch Reading Ease Score) or by a readability measure which has been developed specifically for Navy enlisted personnel (the Kincaid formula).

8.2 The extent to which thinking or conceptual processes are involved

in answering the study questions, which is referred to as Conceptual Level (CL), is related significantly to study question error rate.

8.3 Conceptual Levels can be determined reliably and quickly by judges who are unfamiliar with NRCC content or development.

8.4 Poorly written materials (either study questions or associated MTI passages) may be identified readily as those materials which are written at a low CL but which have high error rates.

8.5 A valid determination could not be made of the extent to which (a) these NRCC materials (for the MT3&2 course) were representative of other NRCCs, or (b) the answers of this sample of MTs in general would be typical of other Navy ratings. These findings do indicate that MTs appear to have significantly lower GCT scores than similar rate levels in other ratings, and have spent fewer months in rate and fewer years in the Navy before taking the advancement exams for comparable rate levels.

9.0 RECOMMENDATIONS. The following recommendations are made in association with the above conclusions:

9.1 The present findings should be replicated using different NRCCs, rates, and ratings. In addition, the 4-point Conceptual Level rating scale should be applied by different judges to these other NRCC materials in order to determine the reliability of the CL rating technique.

9.2 Readability formulas should be evaluated which include some measure of the thinking required to understand and use textual material. A readability formula which should be evaluated in this manner is the

Reading Scale.

9.3 A determination should be made of the relationship between verbal intelligence (as measured by GCT) and performance (error rates) on high and low CL study questions. Such a determination would permit concurrent validation of the CL techniques, as well as provide information about reading and other NRCC aptitudes and skills which could perhaps be improved through education and training.

9.4 More should be done to analyze long-term retention of the information presented in high and low CL study questions, as well as the incidental learning which may occur in association with answering high and low CL study questions. These results would show whether more or fewer high CL study questions should be included in NRCCs.

9.5 The training effectiveness of organizing study questions (and associated RTM material) in a progressive, systematic fashion from low to high CL ratings in relation to instructional objectives should be determined. Such an organization may result in easier learning and better retention of high CL material than would be true for the present NRCCs in which much of the material is presented unsystematically.

9.6 Research should be conducted into using the CL rating technique (or a similar technique) for judging the conceptual difficulty or comprehensibility of graphic material. The information resulting from this research would indicate the aptitude and skill level (and associated training) required to understand and use graphic materials.

9.7 Conceptual Level judgments should be used to correct or discard faulty, ambiguous, misleading, or deceptive ("tricky") study questions and RTM passages from NRCCs. By identifying this material through low

CI study questions which are associated with high error rates, more objective and improved quality control would be exercised over UPCC development, and better training should result.

10.0 REFERENCES

1. Biersner, R. J. Reading grade levels of Navy rate training manuals and non-resident career courses. CNETS Report 2-75, Chief of Naval Education and Training Support, Pensacola, Florida 32509, May 1975.

2. Carver, R. P. Measuring the reading ability levels of Navy personnel. AIR Technical Report, American Institutes for Research, 8555 Sixteenth Street, Silver Spring, Maryland 20910, October 1973.

3. Duffy, T. M., Carter, J. D., Fletcher, J. D., and Aiken, E. G. Language skills: A prospectus for the Naval Service. NPRDC Special Report 76-3, Naval Personnel Research and Development Center, San Diego, California 92152, October 1975.

4. Kincaid, J. P., Fishburne, R. P., Rogers, R. L., and Chissom, B. S. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. CNTECHTRA Research Branch Report 8-75, Chief of Naval Technical Training, Naval Air Station, Millington, Tennessee 38054, February 1975.

5. Powers, R. D., Sumner, W. A., and Kearl, B. E. A recalculation of four adult readability formulas. Journal of Educational Psychology, 49, 99-105, 1958.

6. Smith, E. A., and Kincaid, P. Derivation and validation of the Automated Readability Index for use with technical materials. Human Factors, 12, 457-464, 1970.

7. Flesch, R. Marks of Readable Style: A Study in Adult Education. New York: Bureau of Publications, Teachers College, Columbia University, 1943.

8. Klare, G. R. A Manual for Readable Writing. Glen Burnie, Maryland: REM Co., 1975.

9. Carver, R. P. Improving reading comprehension: Measuring readability. AIR Final Report, American Institutes for Research, 8555 Sixteenth Street, Silver Spring, Maryland 20910, 1974.

10. Long, J. L. Naval correspondence courses: They look like educational materials. CNET SUPPORT Reserve Unit 423 Report, Chief of Naval Education and Training Support Reserve Unit 423, Pensacola, Florida 32509, April 1976.

11. Flesch, R. How to Test Readability. New York: Harper and Brothers, 1951.

12. Bloom, B. S. (Ed.) Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain. New York: David McKay Co., Inc., 1956.

13. Cochran, W. G., and Cox, G. M. Experimental Designs. New York: Wiley, 1966.

14. Rickards, J. P. Interaction of position and conceptual level of adjunct questions on immediate and delayed retention of text. Journal of Educational Psychology, in press.

11.0 FOOTNOTES

1. Lieutenant Commander Biersner was Programs/Systems Effectiveness Evaluation Analyst (Code 01A2) at the Naval Education and Training Support Command, and is now Research Projects Officer (Code 01), Naval Submarine Medical Research Laboratory, Naval Submarine Base New London, Groton, Connecticut 06340. Mr. Doucette is a Mathematical Statistician at the Naval Education and Training Program Development Center, Pensacola, Florida 32509.

2. The subject group varied by plus or minus two subjects from this figure (N = 30) depending on whether or not a subject failed to submit one or more of the twelve assignments described below in section 5.2.

3. The characteristics of a well written NRCC are (a) the written material is presented in a grammatically correct, organized, and concise fashion, (b) the information is technically accurate, and (c) the study questions are judged to be fair and valid (face validity).

4. Section 6.1 provides evidence that this judgement was not entirely valid.

5. A Kincaid RGL was not computed for the HTM 3 & 2 assignment booklet because the Kincaid formula was developed only from sample passages selected from RTMs (see reference 1 for details).

6. These parameters were (a) less than five or more than 37 words per sentence, (b) less than 120 or more than 200 syllables per 100 words, and (c) a Flesch Reading Ease Score of less than 0 or more than 100.

TABLE 1

Summary Statistics of Verbal Intelligence and Career Data for
Navy Personnel Taking Advancement in Rate Examinations

	HT Apprentice	Overall Apprentice	Significance Test	HT 3	Overall PO 3	Significance Test
Average GCT	(N = 531*) 51.40 (sd = 8.58)	(N = 14,663*) 52.90 (sd = 9.34)	t = 3.94 (df = 15,212; p<.001)	(N = 419*) 53.07 (sd = 8.39)	(N = 18,806*) 55.67 (sd = 10.05)	t = 6.09 (df = 19,223; p<.001)
Months in Rate % 12 months or less % over 12 months	(N = 617) 77.31% 22.69%	(N = 16,916) 71.78% 28.22%	χ^2 = 9.011 (df = 1; p<.01)	(N = 510) 52.35% 47.65%	(N = 21,498) 39.27% 60.73%	χ^2 = 30.255 (df = 1; p<.001)
Length of Service % 36 months or less % over 36 months	(N = 617) 85.58% 14.42%	(N = 16,916) 84.47% 15.53%	χ^2 = 0.555 (df = 1; p = ns)	(N = 510) 56.63% 41.37%	(N = 21,498) 46.45% 53.55%	χ^2 = 28.695 (df = 1; p<.001)

*Ns may vary depending on availability of data from CHNAVPERS.

TABLE 2

Summary Statistics of Flesch and Kincaid Reading Grade Levels (RGLs)
for RTM Passages and Study Questions in the Error and Nonerror Samples

	RTM Passages									Study Questions								
	Error Sample				Nonerror Sample					Error Sample				Nonerror Sample				
	N	$\bar{X}$	SD	RGL	N	$\bar{X}$	SD	RGL	t - test*	N	$\bar{X}$	SD	RGL	N	$\bar{X}$	SD	RGL	t - test*
Flesch Reading Ease Score	80	58.12	16.18	10th-12th	65	54.34	17.34	10th-12th	1.31 (p = ns)	80	59.72	18.77	10th-12th	65	55.64	17.46	10th-12th	1.34 (p = ns)
Kincaid RGL	80	9.86	3.29	--	65	10.16	3.31	--	0.54 (p = ns)	80	9.56	3.71	--	65	10.26	4.58	--	1.14 (p = ns)

*Degrees of freedom are 143 for comparisons made in this table.

APPENDIX A

The 4-Point Rating Scale Used to Judge the Conceptual Level (CL) of Study Questions

CONCEPTUAL LEVEL 1 (CL = 1)

The answer and/or question stem are taken verbatim or nearly verbatim from the manual. Even if the meaning of some of the words is not totally understood, the answer is obvious because the words and sentences in the study question nearly duplicate the words and sentences in the manual. Included under this rating would be those cases in which some of the words from the manual may have been removed from the study question (or vice versa), but the key words remain in the same order (rearrangement of key words is therefore unnecessary).

CONCEPTUAL LEVEL 2 (CL = 2)

Simple reordering or rearrangement of some of the words and sentences in the manual is necessary to answer the study question. The answer and/or question stem paraphrases the wording in the manual. For example, numbers presented in graphic form in the manual are translated or converted to verbal form in the study question (or vice versa), or a positive statement in the manual is reversed to a negative statement in the study question (or vice versa). Most of the thinking is involved in rewording, reordering, or reversing the sentences and words in the manual; after this process is complete, little thinking is required to answer the study question.

CONCEPTUAL LEVEL 3 (CL = 3)

A simple situation or problem described in the study question requires the application of one or more general procedures or concepts presented in the manual. All of the information necessary to answer the study question is provided in the manual or the stem of the question, and the interrelationships between steps in the procedures or between concepts have been fully specified. Some care, however, may still be required in following or applying these procedures or concepts. For example, the answers may be similar, and may vary only slightly according to minor deviations in following or applying the procedures or concepts described in the manual; or the information required to answer the study question may be located in several widely separated sentences in the manual, and the reader must organize the sentences together in order to answer the study question. A high level of care and discrimination is therefore required in answering the study question. Although the procedures may be straightforward, some thinking is required in order to insure that the information or conditions which are specified in the study question are applied in the exact sequence or fashion that is indicated in the manual. Moderate levels of thinking are required in converting, translating, and organizing the specific conditions or information provided in question stem and answer to the general procedure or concept presented in the manual. The reader is usually certain of the correct answer after rechecking the steps or logic that have been prescribed in the manual.

CONCEPTUAL LEVEL 4 (CL = 4)

Answering this type of study question also requires the application of one or more general concepts or procedures which are presented in the manual. Some of the relationships between the concepts or steps of a procedure are, however, either vaguely described or remain unspecified in the manual. Much of the interpretation, integration, and organization of the concepts or procedures provided in the manual must therefore be done by the reader in order to understand and answer the study question. The reader may also have to make assumptions about the relationship between information provided in the manual and information provided in the study questions. For example, the study question does not present all of the information necessary to determine which of the answers most fully meets the conditions presented in the manual or, vice versa, the study question contains information or conditions which are not specified in the manual. The reader is often uncertain of the correct answer because of the assumptions that must be made in interpreting concepts and procedures in the manual, or in applying these concepts and procedures to the conditions specified in the study questions.

APPENDIX B

Examples of Study Questions and Associated PTU Passages Rated at Each of the 4 Conceptual Levels (CLs) by the 2 Judges

CONCEPTUAL LEVEL 1 (CL = 1)

Study Question:

The standard taper of pipe threads per foot is

1. 1/4 in.
2. 1/2 in.
3. 3/4 in.
4. 1 in.

Associated RTM Passage:

The NPT, which formerly stood for National Pipe Thread, is still used as a carryover and now refers to the new name for the same thread, American Standard Taper Pipe Threads. The standard taper of pipe threads is 3/4 inch per foot. The number of threads per inch varies according to the size of pipe as follows: 1/16- and 1/8-inch pipe have 27 threads per inch; 1/4- and 3/8 inch pipe have 18 threads per inch; 1/2- and 3/4-inch pipe have 14 threads per inch; 1-, 1 1/4-, 1/2-, and 2-inch pipe have 11 1/2 threads per inch; 2 1/2-inch pipe and pipe larger than 2 1/2 inches have 8 threads per inch.

CONCEPTUAL LEVEL 2 (CL = 2)

Study Question:

The gaskets used between a grinding wheel and it retaining washers should be made of

1. iron or steel
2. fiber or rubber
3. paper or cardboard
4. copper or brass

Associated RTM Passage:

The following operating instructions and safety precautions are applicable in general to all grinders and specifically to the bench and pedestal types.

1. When starting a grinder, push the start button and stand to one side for at least one minute while the machine comes up to full speed. There is always a possibility that a wheel may fly to pieces when coming up to speed.

2. Never force work against a cold wheel; apply work gradually to give the wheel an opportunity to warm. You thereby minimize the possibility of breakage.

3. Handle wheels carefully. Before replacing a wheel on a grinder, always sound the new wheel for cracks. To sound a wheel, tap it lightly with a piece of hard wood; a good wheel gives out a clear ringing sound when tapped, but if the wheel is cracked, a dull thud is heard. Make sure that a fiber or rubber gasket is in place between each side of the wheel and its retaining washer (spindle wheel flange). Tighten the spindle nut just enough to hold the wheel firmly; if the nut is tightened too much the clamping strain may damage the wheel.

CONCEPTUAL LEVEL 3 (CL = 3)

Study Question:

After the frames are positioned and first secured to the keel, they are bolted or riveted to clamp them between the

1. keelson and keel
2. ribbands and keel
3. planking and keel
4. garboard strake and keel

Associated MTM Passage:

When a boat shop is ordered to build a boat the following plans are furnished: lines and off-sets, inboard and outboard profiles, amidship sections, and a data sheet which gives the scantling or sizes of the principal members of the framing. With this information, the boatbuilders proceed to cut out and shape the various parts of the boat.

The first step in assembling a boat is to place the keel on blocks which are heavy enough to support the weight of the boat and high enough to allow for working room. Then the stem and sternpost are riveted or bolted to the keel and held rigidly in position by temporary shores and braces.

Next, the ends of each pair of frames are butted and secured by thwartship timbers known as FLOORS. Then the frames are put in position and secured to the keel. Filling pieces are then fitted over the keel and between the frames. Next the keelson is placed over the frames, floors, and filling pieces and directly above the keel. Bolts or rivets are used through the keelson, filling pieces, and keel, to clamp and hold the frames securely in place. The portion or end of the frame set to keel is known as heel. The other end to top sides is known as head. After frames are positioned and secured to the keel, they are lined up in the vertical direction and secured with temporary wooden battens of heavy sections to fix them rigidly in final position. These battens, known to the boatbuilder as rib-bands, are installed approximately horizontal. Any inside horizontal framing required in the hull may be installed on forms prior to setting transverse framing, or after frames are removed. Hull design dictates which method should be followed.

CONCEPTUAL LEVEL 4 (CL = 4)

Study Question:

Having just replaced a depleted cylinder of oxygen in a portable oxyacetylene outfit, you reconnect the hose and torch. What are the steps of procedure for purging air from the hose and torch, assuring that the oxygen (A) cylinder valve, (B) torch valve, and (C)regulator are closed?

1. Open A, B, and C in order, then close C.
2. Open A, C, and B in order, then close A.
3. Open A, C, and B in order, then close B.
4. Open A, B, and C in order, then close A and B.

Associated M.I. Passage:

5. BACKFIRE AND FLASHBACK: Unless the system is thoroughly purged of air and all connections in the system are tight before the torch is ignited, the flame is likely to burn inside the torch instead of outside the tip. The difference between the two terms backfire and flashback is this: in a backfire, there is a momentary burning back of the flame into the torch tip; in a flashback, the flame burns in or beyond the torch mixing chamber. A backfire is characterized by a loud snap or pop as the flame goes out. A flashback is usually accompanied by a hissing or squealing sound. At the same time, the flame at the tip becomes smoky and sharp pointed. When a flashback occurs, immediately shut off the torch oxygen valve, then close the acetylene valve. The occurrence of a flashback indicates that something is radically wrong either with the torch or with the manner of handling it. A backfire is less serious. Usually the flame can be relighted without difficulty. If backfiring continues whenever the torch is relighted check for these causes: overheated tip, gas working pressures greater than that recommended for the tip size being used, loose tip, or dirt on the torch tip seat. These same difficulties may be the cause of a flashback, except that the difficulty is present to a greater degree. For example, the torch head may be distorted or cracked.

In most instances, backfires and flashbacks result from carelessness. These difficulties can be avoided by making certain that: (1) all connections in the system are clean and tight; (2) torch valves are closed (not open or merely loosely closed) when the equipment is stowed; (3) the oxygen and acetylene working pressures used are those recommended for the torch employed; and (4) the system is purged of air before the apparatus is used. Purging the system of air is especially necessary when hose and torch have been newly connected or a new cylinder is incorporated into the system. Purging a system is accomplished as follows: Close torch valves tightly, then slowly open the cylinder valves. Next, open the regulator slightly. Open the torch acetylene valve and allow acetylene to escape for 5 to 15 seconds, depending on the length of the hose. Close the acetylene valve. Repeat the procedure on the oxygen side of the system. After purging air from the system, light the torch as described previously.

DISTRIBUTION LIST 12-76

Assistant Secretary of the Navy
(Manpower & Reserve Affairs) (2)
Navy Department
Washington, DC 20350

Chief of Naval Operations
(OP-39, 59, 099(2), 964, 987E only)
Navy Department
Washington, DC 20350

Vice Chief of Naval Operations
Navy Department
Washington, DC 20350

Chief of Naval Personnel
(Pers-Od, 5, 6, 2x, 212, 52, 55, 61 only)
Department of the Navy
Washington, DC 20370

Chief of Naval Research
(Code 450 (4), 458 (2) only)
800 North Quincy Street
Arlington, VA 22217

Chief of Naval Education and Training
(CNET N2, 4, 5, 5a only)
Naval Air Station
Pensacola, FL 32508

Chief of Naval Material
(NAVMAT 03T2)
Navy Department
Washington, DC 20360

Commander
Naval Recruiting Command
4015 Wilson Boulevard
Arlington, VA 22203

Chief of Naval Air Training
Naval Air Station
Corpus Christi, TX 78419

Chief of Naval Technical Training (2)
Naval Air Station, Memphis (75)
Millington, TN 38054
(ATTN: Dr. Norman Kerr)

Office of Assistant Secretary of Defense (M&RA)
The Pentagon
Washington, DC 20301

Commander Training Command
U. S. Pacific Fleet
San Diego, CA 92147

Commander Training Command
U. S. Atlantic Fleet
Norfolk, VA 23511

Commander (2)
Naval Training Center
Great Lakes, IL 60088

Commander
Naval Electronics Laboratory Center
San Diego, CA 92152

Commanding Officer
Manpower and Material Analysis
Center, Pacific
San Diego, CA 92132

Commanding Officer
Naval Health Research Center
San Diego, CA 92152

Commanding Officer
Naval Aerospace Medical Institute
Naval Aerospace Medical Center
Pensacola, FL

Commanding Officer
Naval Education and Training Program
Development Center, Ellyson
Pensacola, FL 32509

Commanding Officer
Naval Submarine Medical Center
Naval Submarine Base, New London
Groton, CT 06340

Commanding Officer
Naval Medical Research Institute
National Naval Medical Center
Bethesda, MD 20014

Naval Personnel Research and Development Center
San Diego, CA 92152

Commanding Officer
Service School Command
Building 300
Naval Training Center
Great Lakes, IL 60088

Commanding Officer
Service School Command
Naval Training Center
Orlando, FL 32813

Commanding Officer
Service School Command
Naval Training Center
San Diego, CA 92133

Commanding Officer
Naval Training Equipment Center
Orlando, FL 32813

Commanding Officer
Naval Education and Training Support
Center, Pacific
Fleet Station Post Office Building
San Diego, CA 92132

Commanding Officer
Naval Education and Training Support
Center, Atlantic
Building Z86, Naval Station
Norfolk, VA 23511

Director
Naval Instructional Technology Development Center
Naval Training Center
San Diego, CA 92133

Director
Training Analysis and Evaluation Group
Naval Training Center
Orlando, FL 32813

Human Resources Development Division
U. S. Army Personnel and Administration
Combat Development Activity
Fort Benjamin Harrison, IN 46216

Army Research Institute for Behavioral and
Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

Personnel Research Division
Air Force Human Resources Laboratory (AFSC)
Lackland Air Force Base, TX 78236

Occupational Research Division
Air Force Human Resources Laboratory (AFSC)
Lackland Air Force Base, TX 78236

Commandant of the Marine Corps
(Code MPI)
Navy Department
Washington, DC 20380

Commandant
U. S. Coast Guard
(G-P-1/62 J. Cowan)
Washington, DC 20590

Superintendent
Naval Academy
Annapolis, MD 21402

Superintendent
U. S. Air Force Academy
Colorado Springs, Colorado 80230

Superintendent
Naval Postgraduate School
Monterey, CA 93940

Superintendent
U. S. Coast Guard Academy
Groton, CT 06340

Center for Naval Analyses
1401 Wilson Boulevard
Arlington, VA 22209

U. S. Army Enlisted Evaluation Center
Librarian (PCEC-A)
Fort Benjamin Harrison
Indianapolis, IN 46249

National Science Foundation
Office of Science Information Service
Washington, DC 20550

Defense Documentation Center
Cameron Station, Building 5
Alexandria, VA 22314

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CNETS REPORT-3-76	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* The Relationship of Readability and Conceptual Level to Performance on a Navy Non-Resident Career Course		5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Robert J. Biersner Robert E. Doucette		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Chief of Naval Education and Training Support, Pensacola, FL 32509		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE October 1976 13. NUMBER OF PAGES 45
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Conceptual Level
Correspondence Course
Non-Resident Instruction
Readability
Reading Grade Level
Simple to Complex Development

3 and 2,

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

The relationships of readability (as measured by the Flesch and Kincaid formulas) and conceptual level (CL) to performance on study questions contained in the Non-Resident Career Course (NRCC) Hull Maintenance Technician ~~3&2,~~ NAVEDTRA 91510 were analyzed. Two samples of study questions, a nonerror sample and an error sample, were drawn from this NRCC. The results show that readability between the nonerror and the error →

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

391 660

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

samples, or between the textual passages associated with these samples, does not differ significantly. The results do show, however, that the conceptual level (thinking) required to answer the study questions is related to error rate. That is, high error rate study questions are associated with higher conceptual levels than are zero error rate study questions. Methods are presented for using high and low CL study questions to improve the training effectiveness of NRCCs. The recommendation is made that retention of high and low conceptual level study questions be analyzed because previous research has indicated that better learning and memory may be associated with study questions that are highly conceptual. The recommendation is also made that the relationship between NRCC performance and verbal intelligence (measured through the General Classification Test) be determined because verbal intelligence is most likely involved in performing well on highly conceptual study questions. In addition, readability formulas should be developed to include conceptual level as a factor in determining reading grade levels. The inclusion of a conceptual level factor in readability formulas could improve the validity of these formulas.